Joint Optimization of Transport Cost and Reconstruction for Spatially-Localized Compressed Sensing in Multi-Hop Sensor Networks

Sungwon Lee and Antonio Ortega

University of Southern California

16 December 2010
Table of Contents

1 Introduction

2 Related work
   - Joint routing and compression
   - Compressed Sensing (CS)

3 Problem formulation

4 Proposed approach
   - Low-cost sparse projection based on clustering
   - Maximum energy overlap
   - Centralized iterative clustering algorithm

5 Simulation result

6 Conclusion
Correlated data gathering

- Continuous data gathering using a wireless sensor network
- Spatially distributed data has spatio-temporal correlations
- Compression is required for energy efficiency and longevity
- Many joint routing and compression techniques are proposed

Joint Routing and Compression

- **Transform-based approaches**
  - Wavelet-based approaches \(^4\) \(^5\) \(^6\) and distributed KLT \(^7\)
  - Exploit spatial correlation to reduce the number of bits to be transmitted to the sink

- **Critically sampled approaches**
  - ⇒ cost of gathering scales up with the number of sensors
  - ⇒ undesirable for large deployments of sensors

---


Compressed Sensing approach

- Requires smaller number of measurements than dimension of the signal
- Number of measurements depends on the characteristics (sparseness)
- Most computations take place at the sink rather than sensors

Existing approaches on

- number of measurement
- sparseness of measurement system

rather than total transport cost

Our previous work a *heuristic* way to achieve energy saving using spatially-localized measurement system.  

---


Compressed Sensing Basics 11 12 13

- CS addresses three inherent inefficiencies of transform-based approaches
  - Large number of data samples
  - Compute all transform coefficients
  - Locations of large coefficients

- Compressed representation by dimension reduction
  \[ y_{M \times 1} = \Phi x_{N \times 1}, \quad M < N \]

- Recover a signal, \( x \in \mathbb{R}^N \), from the representation
  \[ \Rightarrow \text{Recover } x \text{ subject to } y = \Phi x \]
  \[ \Rightarrow \text{Unique solution in under-determined problem?} \]

---

- Assume that a signal, $x \in \mathbb{R}^N$, is $k$-sparse in a given basis: $\Psi x = \Psi a, |a|_0 = k$, where $k \ll N$
- Replace data samples with few linear projections, $y = \Phi x$.

Mutual Coherence: $\mu = \max_{i,j} |H(i,j)|$, where $H = \Phi \Psi$.

Exact recon. when $M = O(K \log N)$ and $\mu$ is small (incoherent).
  - Dense Random Projection (DRP), $\Phi(i,j) \sim N(0, 1)$, works universally.
  - Solve convex unconstrained optimization problem
    $\min_a \frac{1}{2} \|y - \mathbf{H}a\|_2^2 + \gamma \|a\|_1$, where $\mathbf{H} = \Phi \Psi$
Problem Formulation

- **Goal**: Minimize transport cost of measurements with less degradation.
  - Low-cost sparse $\Phi$
  - Proper $\Phi$ with a given basis, $\Psi$, for successful reconstruction
- **Link between CS measurements and data aggregation**

![Diagram](image_url)

(a) Aggregation  
(b) Measurement matrix

- Measurements in CS $\iff$ aggregates on WSN
- Sparsity of $\Phi$ $\iff$ no. of active sensors for aggregation.
- Position of non-zero entries in $\Phi$ $\iff$ aggregation strategy.
Problem formulation

- Traditional CS assumes *DENSE* random $\Phi$
  $\Rightarrow$ every sensor transmits its data once for each measurement
  $\Rightarrow$ high energy consumption (higher than raw data transmission)

- Sparse CS
  $\Phi$ contains a few non-zero entries
  - *SPARSE* random $\Phi$ without considering positions of active sensors \(^{14}\)
  - *SPARSE* measurements on SPT but reconstruction quality is not satisfactory \(^{15}\)

- To reduce transmission cost,
  - Reduce the number of samples for each measurement
  - Aggregate samples of sensors close to each other

  $\Rightarrow$ spatially-localized sparse projection is required


Proposed approach

Low-cost sparse projection based on clustering

- Spatially-localized projection based on clustering
  1. Divide network into clusters of adjacent sensors
  2. Force projections obtained only from all the nodes in each cluster
  3. Localized measurements are transmitted along SPT to the sink

- Example
  Four sensors ($S_1 \sim S_4$) with two clusters ($\{S_1, S_3\}$ and $\{S_2, S_4\}$)

  - Two data aggregation: one for each cluster
    1. $S_3 \rightarrow S_1 \rightarrow Sink$
    2. $S_4 \rightarrow S_2 \rightarrow Sink$

  - Generate measurements from clusters
    1. $y_1 = w_{11}x_1 + w_{31}x_3$, where data sample $x_i$ of $S_i$
    2. $y_2 = w_{21}x_2 + w_{41}x_4$, where r.v. $w_{ij}$ for $S_i$

---

Proposed approach

Low-cost sparse projection based on clustering

Matrix formulation

\[
\begin{bmatrix}
  y_1 \\
  y_2 
\end{bmatrix} =
\begin{bmatrix}
w_{11} & w_{31} & 0 & 0 \\
0 & 0 & w_{21} & w_{41}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_3 \\
x_2 \\
x_4
\end{bmatrix}
\]

\[
= \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
w_{11} & w_{31} & 0 & 0 \\
w_{12} & w_{32} & 0 & 0 \\
0 & 0 & w_{21} & w_{41} \\
0 & 0 & w_{22} & w_{42}
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4
\end{bmatrix}
\]

Leads to Sparse block-diagonal matrix and Clustering matrix

Similar to recent work\(^{17, 18}\) proposed for fast CS computation

Showed comparable results to dense random projection matrices

What is the best clustering matrix ?

⇔ What is the best aggregation strategy for CS reconstruction ?

---


Maximum energy overlap

- Not obvious how localized gathering impacts reconstruction quality
- How to evaluate CS reconstruction accuracy w.r.t. clustering matrix? ⇒ Maximum energy overlap ($\beta$) between clusters and basis functions
- Example of 4 clusters and 3 basis functions

(a) overlap in spatial domain (b) overlap in sparsifying basis

⇒ $\beta = 1$ because of overlap between $C_1$ and $B_1$ (normalized $B_i$)
Proposed approach

Maximum energy overlap

Definition

Maximum energy overlap, $\beta(\Psi)$

$$
\beta(\Psi) = \beta(\tilde{P} \tilde{\Psi}) = \max_{i,j} \sum_l \Psi_i^2(l, j), \quad \beta(\Psi) \in [0, 1]
$$

- $\beta$ shows energy of basis functions captured by each cluster
- Intuitively, measurements from clusters overlapped with more basis functions can convey more information
- $\beta$ decreases
  - $\Rightarrow$ More evenly distributed energy over overlapped clusters
  - $\Rightarrow$ Higher chance for successful reconstruction

---

Theorem

For a given signal $\mathbf{x} = \Psi \mathbf{a}$ with $|\mathbf{a}|_0 = K$ and a clustering (permutation) scheme with parameter $\beta \in [0, 1]$, the $l_1$ optimizer can recover $\mathbf{x}$ exactly with high probability if the number of measurements $M$ satisfies

$$M = O(K\beta N_c \log^2 N)$$

- Assumes that non-overlapped $N_c$ clusters with equal size and $\Phi(i, j) \sim N(0, N_c/N)$
- Sketch of proof
  1. Derive asymptotic upper bound on $\mu$ using Bernstein inequality \(^{20}\)
  2. Derive minimum $M$ using Candes’s result \(^{21}\)
- Smaller $M$ can be achieved by reducing $\beta$

Centralized iterative clustering algorithm

- How to jointly optimize transport cost and reconstruction accuracy?
- $D$ increases (transport cost per measurement increases) \Rightarrow $\beta$ is likely to decrease (smaller number of measurements)
- Design a centralized iterative clustering algorithm
  1. Choose $N_c$ nodes; one for each cluster
  2. For each cluster, find an edge to minimize weight, $W(e) = D(e) + \lambda \beta(e)$
  3. Update $W$
  4. Repeat until every node is assigned to one of $N_c$ clusters
- Similar to Prim’s algorithm to find minimum spanning tree (MST)
- Additional constraints
  - Construct $N_c$ clusters instead of a tree
  - Edge weight can change every step
Reconstruction accuracy and $\beta$

- **Environment**
  - 1000 synthesized sparse data
  - $N_c = 16$ for 20 different clustering scheme
  - Gradient Pursuit for Sparse Reconstruction (GPSR) \(^{22}\)

- **Evaluation**
  - Data is perfectly reconstructed if $\max |x - \hat{x}| < 10^{-3}$
  - Compute Pearson’s linear correlation coefficient, $r \in [-1, 1]$ between
    - $M_{sim}$ is the smallest $M$ that satisfies a perfect reconstruction rate larger than 0.99.
    - $M_{est}$ is computed by Theorem

<table>
<thead>
<tr>
<th>$N = 1024, N_c = 16$</th>
<th>$DB_4$</th>
<th>$DB_6$</th>
<th>$DB_8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>K=20</td>
<td>0.468</td>
<td>0.429</td>
<td>0.513</td>
</tr>
<tr>
<td>K=38</td>
<td>0.567</td>
<td>0.503</td>
<td>0.621</td>
</tr>
<tr>
<td>K=55</td>
<td>0.694</td>
<td>0.589</td>
<td>0.617</td>
</tr>
</tbody>
</table>

Joint Optimization

- $K = 38$ in 2D Daubechies-4 basis with the $2^{nd}$ level of decomposition
- Transport cost = $\sum (\text{bit}) \times (\text{number of hops})$
- $\beta$ vs. average number of hops per measurement. Each point corresponds to the result of the algorithm with different $\lambda$. 

![Graph showing Beta vs. Average hop distance(D) per measurement(M)]
Performance Comparison

- Transport cost ratio vs. MSE
- Compare joint optimization with two different $\lambda$’s with $SPT_{64}$; $SPT_{64}$ showed the best performance of other existing CS approaches. 

---

Conclusion

- Proposed a framework using spatially-localized compressed sensing
- Introduced maximum energy overlap ($\beta$) to estimate reconstruction accuracy
- Proved a relationship between $\beta$ and minimum $M$ for perfect reconstruction
- Proposed a centralized iterative algorithm for joint optimization
- Joint optimization shows better performance than existing CS methods
Thanks!